Respondents will discuss with USEPA the need to conduct bail-down tests in order to determine if residual NAPL is bound to the soil or if NAPL that may be present at or beneath the water table.

### 5.7.2 DP&L Groundwater Investigation

Concurrent with the OU1 groundwater investigation described in the preceding section, the Respondents propose to complete VAS along the northern DP&L property boundary in order to determine if the source of deep groundwater contamination is related to DP&L site activities or potentially caused by off-Site sources (e.g., from Delphi operations). The Respondents propose to first complete VAS investigation at the proposed location between MW-222A and MW-224B in order to evaluate aquifer data in the vicinity (Figure 5.5). Should the analytical results from this proposed VAS location indicate no deep groundwater contamination, the Respondents propose to complete VAS investigation at the proposed location east of MW-224B in order to determine whether the contamination is present further to the east. The Respondents propose to complete the VAS location(s) at the north end of the DP&L property to a depth of 150 ft bgs as this is approximately the lower limit of cis-1,2-DCE contamination and the core of the vinyl chloride contamination identified in VAS borings immediately to the south and should provide sufficient indication as to whether cis-1,2-DCE and vinyl chloride are present in the Lower Aquifer Zone beneath this area.

The Respondents also propose to complete a VAS boring immediately south of the DP&L Service Building between the locations of (BH11-13 DP&L and BH12-13 DP&L) (Figure 5.5). The intent of the VAS boring is to determine whether the low concentrations of TCE identified in DP&L boreholes BH11-13 and BH12-13 during the DP&L investigation are indicative of greater TCE contamination in deeper groundwater and to further delineate the deeper cis-1,2-DCE contamination present in the Lower Aquifer Zone in this area. The Respondents propose to complete the VAS boring south of the Service Building to a depth of 100 ft bgs as this should adequately confirm whether shallow TCE impacts are present and identify the peak cis-1,2-DCE concentrations in the Lower Aquifer Zone. The proposed DP&L VAS locations are presented on Figure 5.5. The need for additional investigation (i.e., installation of permanent monitoring wells and the appropriate screened intervals) will be determined based on the results of the VAS at these locations.

### 5.7.3 OU2 Groundwater Investigation

GHD will propose the scope of an OU2 Groundwater Investigation following completion of the groundwater investigations described in Sections 5.7.1 and 5.7.2. The Respondents will provide a description and rationale for proposed OU2 groundwater locations to USEPA for review prior to implementing the work. OU2 groundwater locations will be proposed in areas where impacted groundwater may potentially migrate from within OU1, based on OU1 soil/fill and groundwater characterization and groundwater flow direction.

# 6. Background Comparisons

For elements of the investigation requiring a comparison to background conditions, the following methodology will be used. Such comparisons are noted particularly for the following investigation elements, although the methodology presented herein may also be applied to additional items, if identified, during the course of the investigation.

- Site Soil, Phase 1B (Comparison to Background)

The purpose and methods for making comparisons to background may be different than those for attributing contamination to an off-Site source. GHD will factor in the possibility of widespread contamination, naturally-occurring parameters; and ubiquitous anthropogenic contaminants, etc. prior to making a determination.

The constituents that will be included for background comparisons are metals and PAHs for soils. GHD will collect near-surface soil/fill samples from 0 to 2 ft bgs and compare the analytical data to background soil/fill sample data from the same depth for metals and PAHs. This comparison to background is proposed because metals and PAH contamination may be due to regional atmospheric deposition from industrial and coal-burning activity in the broader surrounding area of the Site and any impacts would be expected in the upper layers of soil. GHD will collect subsurface soil/fill samples from 2 to 15 ft bgs and compare the analytical data to background sample data from the same depth for metals only (which may be naturally occurring in local soils).

Figure 6.1 presents proposed background soil sampling locations. Background reference soil sampling locations will be identified in areas outside a reasonable zone of potential influence (via surface runoff or substantial airborne dust deposition) for the Site and based on areas that have had little or no industrial impact (i.e., Carillon Park to the east and the cemetery to the north). The additional background location outlined on Figure 6.1 (i.e., northern part of Parcel 3264) is an area that has no known history of industrial use or dumping. A review of historical aerial photographs compiled in previous investigation reports (CRA, 2010) indicates that Parcel 3264 was undeveloped and possibly used for agriculture from the 1950s to the 1970s, before becoming heavily forested in the late 1980s. There is no visual evidence in the aerial photographs that landfilling or excavation activities from the SSDL to the north or industrial activities from companies to the south encroached onto Parcel 3264. GHD will cease sampling activities if potential contaminants and/or non-native fill material are identified at any of the background sampling locations. The Respondents will submit work plan addenda (letters) to USEPA for each sample medium. The work plan documents will include details on the proposed background sample locations.

## 6.1 Background Comparison Approaches

Evaluation of site vs. background conditions using environmental quality data is typically carried out using either group-based or individual-based statistical comparisons. Group-based comparisons pool the data from a number of samples collected at a site (e.g., from within an area of interest) and contrast these against a pooled set of background samples. In such a case, a determination may be made as to whether or not the site area of interest as a whole is consistent with or above background conditions. In contrast, individual-based comparisons make a decision (i.e., consistent with or above background) for each investigative location at the site. In terms of the different elements of the proposed investigations, group-based background comparisons may be applicable for portions of the baseline risk assessment, but the majority of testing will consider individual point comparisons (site vs. background) for the purposes of identifying and delineating potential areas of the Site that appear to have contaminants present above background conditions.

For individual-based comparisons against background, the statistical approaches employed typically establish an expected range (e.g., 95th or 99th percentile) of contaminant concentrations based on the background sample results, against which the site data are compared. An on-site measurement falling outside of the expected background range is identified as being potentially impacted, and is further evaluated to confirm this finding (e.g., using confirmatory sampling or considering the spatial patterns of results in other site samples collected nearby). Confirmation is

required due to the statistical nature of the background expected range calculations, which result in infrequent occurrence of background conditions outside of the range (e.g., 1 in 20 background samples for a 95th percentile range, or 1 in 100 for a 99th percentile range).

For group-based comparisons against background, the statistical approaches employed typically compare the site and background groups based on distributional characteristics (e.g., mean, median, or percentile values) through the use of hypothesis testing. In carrying out such tests, statistically-significant findings provide strong evidence that contaminant concentrations found in the area of interest on the site are different than those present in background areas.

When designing and implementing an environmental investigation where background comparisons are to be made, it is important to try to match background sampling media to those present at the site, as far as is possible. That is, matching soil types/textures (including multiple soil types if necessary due to site stratigraphy), groundwater aquifers, etc. This prevents the finding of differences between site and background conditions due to factors unrelated to activities at the Site (e.g., different native mineralogy in different soil layers under a site).

## 6.2 Relevant Guidance and References

The issue of appropriate background comparison techniques is discussed in numerous guidance and environmental statistic texts. The methods proposed for the investigations have been selected for consistency with the following documents.

- USEPA, June 1994. Statistical Methods for Evaluating the Attainment of Cleanup Standards. Volume 3: Reference-Based Standards for Soil and Solid Media. Environmental Statistics and Information Division (2163), Office of Policy, Planning, and Evaluation. EPA 230-R-94-004.
- NAVFAC, April 2002. Guidance for Environmental Background Analysis, Volume I: Soil. Naval Facilities Engineering Command, Washington, DC. NFESC User's Guide UG-2049-ENV.
- NAVFAC, April 2003. Guidance for Environmental Background Analysis, Volume II: Sediment. Naval Facilities Engineering Command, Washington, DC. NFESC User's Guide UG-2054-ENV.
- NAVFAC, 2004. Guidance for Environmental Background Analysis. Volume III: Groundwater. Naval Facilities Engineering Command, Port Hueneme, CA. User's Guide UG-2059-ENV.
- USEPA, September 2002. Guidance for Comparing Background and Chemical Concentrations in Soil for CERCLA Sites (OSWER 9285.7-41). Office of Emergency and Remedial Response, United States Environmental Protection Agency, Washington, DC. EPA/540/R-01/003.
- Ohio EPA, July 2014. Evaluation of Background Metal Soil Concentrations in Montgomery County – Dayton Area; Summary Report for Ohio EPA's Voluntary Action Program. Ohio Environmental Protection Agency.
- USEPA, February 2006. Data Quality Assessment: Statistical Methods for Practitioners (EPA QA/G-9S). Office of Environmental Information, United States Environmental Protection Agency, Washington, DC. EPA/240/B-06/003. [Available at http://www.epa.gov/QUALITY/qs-docs/g9s-final.pdf]. [Section 3.3 in particular].
- USEPA, March 2009. Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities – Unified Guidance. Office of Resource Conservation and Recovery, Program Implementation and Information Division, United States Environmental Protection Agency Washington, DC. EPA 530-R-09-007. [Chapter 5 and elsewhere].

- USEPA, September 2013. ProUCL Version 5.0.00 Technical Guide. United States Environmental Protection Agency, Office of Research and Development, Washington, DC. EPA/600/R-07/041. [Chapters 3 and 5].
- USGS, 2002. Statistical Methods in Water Resources. By D.R. Helsel and R.M. Hirsch. Chapter A3 of Book 4, Hydrologic Analysis and Interpretation in Techniques of Water-Resources Investigations of the United States Geological Survey. [Available at http://pubs.usgs.gov/twri/twri4a3/]. [Chapter 3].

For the purposes of individual-based background comparisons (e.g., used in detection monitoring or for delineation of contamination), a general approach found in these references is to use a statistical tolerance or prediction limit to establish a background threshold value (BTV), which is the upper[19] expected range of background concentrations given by a certain percentile of background (e.g., 95th or 99th). Consequently, for elements in the present investigation where individual-based background comparisons are required, BTVs based on statistical upper tolerance limits (UTLs) for the 95th and/or 99th percentile of background have been selected for use. A detailed discussion of UTL calculation methods is found in Chapters 3 and 5 of USEPA's ProUCL version 5.0.00 technical guide (2013, see list above).

For the purposes of group-based background comparisons (e.g., when comparing contaminant concentrations within an area of concern vs. background as part of a risk assessment), different hypothesis tests are available in the references above. Where certain statistical assumptions are met by the data sets considered (e.g., normal distribution, homogeneity of variance), parametric statistical tests are available (e.g., analysis of variance, Student *t*-test). Where these assumptions are not met by the available data, analogous non-parametric (rank-based) statistical methods are available (e.g., Mann-Whitney/Wilcoxon Rank-Sum test, modified Quantile test). Where required for the present investigation, statistical group comparisons will be carried out using the Mann-Whitney/Wilcoxon Rank-Sum test and the modified Quantile test, supplemented by the Student *t*-test where assumptions of the parametric test are met.

Ohio EPA previously conducted a background study of metals concentration in soils in the Dayton area (Ohio EPA, 2014). The chemical analysis results are included in the Ohio EPA report, and may provide additional data to supplement background soil samples collected around the Site (only for the metals reported in the Ohio EPA study). The site-specific background values for metals in soil developed in the RI/FS will also be compared to the Montgomery County background values, to assess whether local conditions are similar to those found regionally.

It is noted that the Ohio EPA (2014) background study utilized similar statistical approaches for individual-based background comparisons, i.e., calculating an upper percentile estimate. In the case of normally-distributed data, an estimate of the 97.5th percentile was used (mean + 2 standard deviations). For all other data sets, a 95$^{th}$ percentile UTL was calculated using USEPA's ProUCL (version 4.1).

[19] In certain cases, a lower limit may also be considered, e.g., for pH or oxygen content in water, but upper limits are much more commonly encountered.

## 6.3 Statistical Considerations

In order to achieve an appropriate and successful statistical comparison of site and background conditions, a number of factors will be considered during sampling design and data analysis. These factors include:

- Background sample size – a minimum of eight to ten background samples will be collected for each environmental medium, as applicable, and/or stratum within the medium (e.g., different soil types and/or aquifers).
- The desired minimum confidence level to be used in the statistical comparisons is 95 percent (i.e., statistical significance of $\alpha = 0.05$).
- The specific statistical method used needs to be appropriate for the observed characteristics of the site and/or background data sets obtained. This requires assessing each data set for the following statistical parameters:
  - Percentage of non-detect values.
  - Statistical data distribution (e.g., testing for normal, gamma and lognormal distributions, per USEPA's ProUCL version 5.0.00 software's approach).
  - Statistical outliers, particularly in background data sets (e.g., testing for outliers via probability plotting and outlier tests).
- QA/QC samples – where field duplicate samples are collected and submitted for laboratory analysis, the resulting data will be averaged prior to statistical calculations in order to avoid over-weighting the sampling location where duplicates were collected. Any laboratory duplicate QA/QC sample results will be excluded from the statistical analyses.
- Confirmatory analysis and/or resampling – for point-based background comparisons using BTVs, it is recognized that periodic occurrence of parameter concentrations above a BTV are expected due to natural variation in the background population (e.g., 1 in 20 samples for a 95th percentile based BTV). Where a site observation exceeds the 95th percentile BTV, it will additionally be compared to a 99th percentile BTV. If the result falls below the 99th percentile BTV, and no spatially- adjacent observations also exceed the 95th percentile BTV, the site observation will be considered to not indicate a site-related effect. However, if the site result exceeds the 99th percentile BTV, or another adjacent site result also is above the 95th percentile BTV, then it will be considered to indicate an above-background condition, unless a confirmatory resample is collected and found to not be above the BTV.

## 6.4 Summary of Statistical Methods Selected for Background Comparisons

In consideration of the information presented above, as well as the objectives of the present investigation as detailed in the DQO tables, the following methods will be used for comparing contaminant concentrations in environmental samples collected at the Site against concentrations observed in background samples.

1. For point-based comparisons (i.e., as described for Phase 1B of the different investigations described in the DQO tables for all media except soil gas), BTVs will be calculated using the available background data:

- If greater than half of the background data are non-detects, if a background data set is not found to follow a discernible statistical distribution, or if one or more statistical outliers are present in a background data set, then a non-parametric UTL on the 95th percentile of background (with 95 percent confidence) will be generated for use as the BTV. This will be done following the methods in USEPA's ProUCL version 5.0.00 software (USEPA, 2013).
- If at least half of the points in a background data set are detects, a discernible statistical distribution (normal, gamma or lognormal) is identified, and no statistical outliers are present, then a parametric UTL on the 95th percentile of background (with 95 percent confidence) will be generated for use as the BTV. This will be done following the methods in USEPA's ProUCL version 5.0.00 software (USEPA, 2013).
- Individual site data will be compared against the BTVs as follows:
  - Where a site observation exceeds the 95th percentile BTV, it will additionally be compared to a 99th percentile BTV.
  - If the result falls below the 99th percentile BTV, and no spatially-adjacent observations also exceed the 95th percentile BTV, the site observation will be considered to not exceed background conditions.
  - However, if the site result exceeds the 99th percentile BTV or another adjacent site result also is above the 95th percentile BTV, then it will be considered to indicate an above-background condition, unless a confirmatory resample is collected and found to not be above the BTV.

To summarize, the following decision matrix will be used for interpreting the results of point-based comparisons against background:

| Comparison Result | Conclusion |
|---|---|
| Below 95th percentile BTV | Not above background |
| Above 99th percentile BTV | Above background |
| Between 95th and 99th percentile BTVs: | |
| - no adjacent sample above 95th percentile BTV | Not above background |
| - adjacent sample above 95th percentile BTV | Above background |

2. For group-based comparisons (i.e., as described for part of the baseline risk assessment):
   - For cases where a particular analyte has not been detected in either background or site samples, no statistical testing will be carried out.
   - If both the site and background data sets contain few non-detects (less than 10 to 15 percent), and follow a normal data distribution, the non-detects will be substituted with a value of one-half their detection limit and the two groups compared using a Student's t-test at 95 percent confidence.
   - In all cases where the site and background data sets combined contain up to 50 percent non-detects, non-parametric testing will be carried out contrasting the two groups using the Mann-Whitney/Wilcoxon Rank-Sum test and the modified Quantile test at 95 percent confidence. For cases where a Student t-test has already been performed, this will be considered as a confirmatory test.

- For the remaining cases where an analyte has been detected in one or more samples, but in less than half of the samples in the pooled site and background data sets, alternate statistical comparisons will be carried out on a case-by-case basis. This could include procedures such as a test of proportions in conjunction with the modified quantile test.

To summarize, the following decision matrix will be used for interpreting the results of group-based comparisons against background:

| Data Characteristics | Test Performed | Result and Conclusion |
|---|---|---|
| No detects | None | Not Above Background |
| ≤15% Non-Detects and Normal Distributions (for each data group – background and Site) | Student's *t*-Test | If at least one test performed finds a statistically significant result having Site > Background: → Above Background |
| ≤50% Non-Detects (for the combined data set – background + Site) | Mann-Whitney/WRS & Quantile Test | If no statistically-significant test results are obtained: → Not Above Background |
| >50% Non-Detects (for the combined data set – background + Site) | Case-By-Case Selection (e.g., Quantile Test & Test of Proportions) | Otherwise: → Not Above Background |

The DQO Table (Tables 5.1 to 5.6) specify whether the Respondents will apply individual-based or group-based comparisons for each study question.

# 7. Baseline Risk Assessment and Ecological Risk Assessment

Major components of the Baseline Risk Assessment (BRA) include constituents of potential concern identification, exposure assessment, toxicity assessment, and human health and ecological risk characterization. The BRA will be prepared for the media in OU1 and OU2.

### *Baseline Human Health Risk Assessment*

GHD proposes to conduct the HHRA (or BRA) in accordance with *Risk Assessment Guidance for Superfund (RAGS Parts A-F)*. These guidance documents, along with the *Exposure Factors Handbook* and *Cancer Risk Assessment* guidelines, are the default guidance documents for risk assessment under CERCLA. There are four key steps to the HHRA process: Data Collection and Evaluation, and Hazard Identification; Exposure Assessment; Toxicity Assessment; and Risk Characterization.

### *Data Collection and Evaluation, and Hazard Identification*

Adequate definition of the Site characteristics and the nature and extent of impacts is an integral component of any risk assessment and is required to reduce uncertainty in the risk assessment findings. The selection of chemicals of potential concern (COPCs) will follow USEPA RAGS Part A, and all chemicals will be screened against the most recent/current applicable USEPA RSLs based on a cancer risk of 1E-06 and hazard quotient of 0.1. Additionally, soil gas and groundwater data will be screened against the most recent USEPA/ODH VISLs. For each medium, chemicals with maximum concentrations less than their respective screening value will not be identified as COPCs,